Disseny, implementació i configuració d'una solució SIEM al CVC

Ferran Ruiz Castellà

Resum—El projecte consta del disseny, la implementació i la configuració d'una solució SIEM al Centre de Visió per Computador. Tot i que el projecte es del àmbit de la ciberseguretat no s'ha limitat només en el SIEM sinó que s'ha enfocat el projecte com a una assessoria de ciberseguretat tractant al CVC com a un client que ha requerit que es desplegui una solució SIEM al seu entorn. El projecte consta d'un anàlisi previ del estat de la ciberseguretat al CVC, el disseny de l'arquitectura de la solució, la definició de casos d'us personalitzats per al CVC, la implementació del aplicatiu tal i com s'ha dissenyat, l'enriquiment del aplicatiu de dades rellevants a nivell de seguretat i un post-anàlisi d'aquests esdeveniments. S'han explotat diverses funcionalitats que el SIEM escollit proporciona. Es pretén que, al finalitzar el treball, els tècnics del CVC disposin d'un SIEM desplegat que els ajudi a orquestrar les seves operacions de seguretat.

Paraules clau—SIEM, Ciberseguretat, Gestió d'esdeveniments, SOC, Centre d'Operacions de Seguretat, Fonts d'informació, Intel·ligència, QRadar, IBM, Casos d'ús, Centre de Visió per Computador.

Abstract—This project consists in the design, implementation and configuration of a SIEM solution in Centre de Visió per computador. This project's scope is cybersecurity. Even if the project's scope is purely cybersecurity it's not limited to the SIEM, we have focused this project as a cybersecurity assesment treating CVC as a client that has requested a SIEM solution to be deployed in its environament. This project contains a previous analysis of the State of cibersecurity at CVC, the design of the solution's arquitecture, the definition of custom use cases, the implementation of the application as it has been designed, the enrichment of the application with relevant security data and the posterior anlisis of that data. Multiple functionalities that the chosen SIEM is able to perform have been exploited. The intention is that, at the end of the project, CVC technicians have at their hand a fully deployed SIEM that helps them orquestrate their security operations.

Index Terms-SIEM, Cybersecurity, Log management, SOC, Security Operations Center, Log sources, Intelligence, QRadar, IBM, Use cases, Centre de Visió per Computador

1 Introducció

En la actualitat, totes les empreses i organitzacions que tinguin una infraestructura de TI són vulnerables a atacs informàtics que poden tenir greus conseqüències com el robatori de informació, la degradació dels actius o l'aturada de processos de negoci. Per poder detectar eficaçment potencials incidents de seguretat es requereix tecnologia que sigui capaç de fer-ho.

Cada dia són més les solucions a disposició dels equips de TI per a fer front al creixent nombre de ciberatacs que es donen: tallafocs d'última generació, anti-virus, sistemes IDS, sistemes IPS, etc. I cada cop són més els actius dintre d'una organització que generen informació rellevant a nivell de seguretat: Controladors de domini, DNS, DHCP, servidors de correu, proxies, etc.

Molts cops els equips de TI es troben abromats per la quantitat d'eines a operar i la quantitat d'actius a monitoritzar i els hi és difícil correlar esdeveniments que succeeixen a diferents parts de la organització (p.e relacionar una connexió sospitosa detectada pel tallafocs amb una infecció d'un endpoint a la mateixa hora).

D'aquesta necessitat neixen les solucions SIEM. Un SIEM (System Information and Event Management) permet gestionar tots els esdeveniments de seguretat que es succeeixen en una organització de manera centralitzada, proporcionant capacitats de detecció, anàlisi i resposta davant d'incidents de seguretat[1].

El SIEM proporciona anàlisi en temps real d'esdeveniments de seguretat generats per aplicacions, eines perimetrals i hardware de xarxa aplicant indexació, regles i intel·ligència per detectar potencials incidents de seguretat utilitzant tant regles personalitzades per a la organització com bases de dades d'intel·ligència globals.

A part de proporcionar anàlisi i detecció a temps real d'incidències de seguretat, un SIEM habilita als equips de TI

- *E-mail de contacte: ferran.ruiz@e-campus.uab.cat*
- *Menció realitzada: Enginyeria del Software*
- *Treball tutoritzat per: Andreu Pérez*
- *Curs 2020/2021*

per realitzar correlació d'esdeveniments, agregació de dades, reporting, gestió de esdeveniments, capacitats forènsiques i gestió sobre les vulnerabilitats de seguretat dels actius de la organització.

El projecte proposat és el disseny, la implementació i la configuració d'una solució SIEM al Centre de Visió per Computador (CVC). Per a poder portar a terme aquesta tasca ha sigut necessari realitzar un estudi del estat actual del CVC a nivell de seguretat, realitzar la implantació del SIEM al seu entorn tecnològic i enriquir al SIEM d'informació per a realitzar una configuració inicial.
Com a part del projecte també es realitzaran, a demanda del CVC, una sèrie de sessions de demostració i formació on s'explicarà als tècnics interessats les capacitats del SIEM, com explotar-les al màxim nivell i com escalar el SIEM conjuntament amb el entorn tecnològic del CVC.

2 Estat de l'art

La ciberseguretat és la pràctica que protegeix sistemes, xarxes i aplicacions d'atacs digitals. Aquests atacs digitals tenen com a objectiu accedir, canviar o destruir informació sensible i pot alterar o interrompre processos de negoci.

En els últims anys, els ciberatacs han augmentat en freqüència i en complexitat. És per això que la única resposta coherent per part d'organitzacions que vulguin protegir els seus sistemes és millorar la qualitat de les seves operacions de seguretat.

Estudis recents[2] determinen que el pressupost pels equips de seguretat ha augmentat fins a un 35000% en els últims 15 anys. Això és una clara indicació de que, a les empreses, cada dia hi ha una preocupació més clara per sostenir la confidencialitat, la integritat i la disponibilitat dels seus sistemes informàtics.

Aquestes preocupacions no són en va, el cibercrim és un negoci que creix cada any i cada cop són més les empreses afectades per atacs informàtics, una breu ullada a algunes estadístiques[3] sobre aquest sector en dels últims anys és suficient per adonar-se de la magnitud del imperi del cibercrim:

- Fugues de informació van exposar 36 trilions de registres a la primera meitat del 2020
- El cost mitjà d'una fuga d'informació per a una empresa es de $3.86M
- El 58% de les fugues d'informació al 2020 contenien dades de caràcter personal
- El preu mitjà de rescat per un atac de ransomware va pujar un 33% al 2020
- Els atacs a dispositius IoT va augmentar un 300% al 2019

Aquestes són només algunes de les dades que ens indiquen que cada cop hi ha més incentius per actors maliciosos per voler atacar a una organització i, com a resposta, és necessari que les organitzacions tinguin un pla definit de governança de ciberseguretat.

Per governar sobre la seguretat d'una empresa o organització cada cop es confia més en un SOC (Centre d'Operacions de Seguretat, per les seves sigles en anglès). Un SOC és una unitat centralitzada dintre d'una organització que empra persones, processos i tecnologia per realitzar monitorització continua i millorar la postura de seguretat de la organització prevenint, detectant, analitzant i responent a incidents de ciberseguretat.

Un SOC ofereix, entre altres, les següents capacitats a una organització

- Monitorització proactiva i continuada dels entorns IT
- Rapida detecció i resposta a incidències de seguretat
- Prevenció de incidents
- Protecció dels actius de la organització
- Disseny i implementació del marc de control de ciberseguretat de la organització
- Auditoria regular dels sistemes de la organització per assegurar el compliment de les normatives actuals

Com hem comentat, un SOC consta de 3 recursos fonamentals: Persones, Processos i Tecnologia[4].

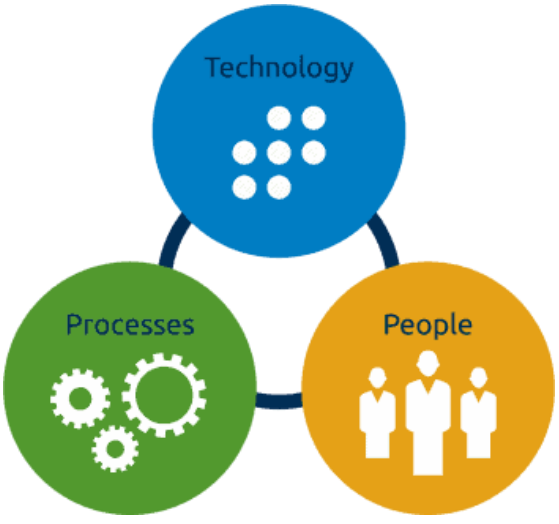


Figura 1. Diagrama SOC

L'abast d'aquest projecte no és muntar un SOC al complert, però si que es pretén millorar les capacitats del CVC a l'hora de detectar i respondre a incidents de seguretat. Per fer això, ens centrarem en la tecnologia i farem també una visita als processos més bàsics a l'hora de realitzar operacions de seguretat.

3 El CVC i el seu estat a nivell de seguretat

El CVC és un consorci públic i sense ànim de lucre que es dedica principalment a la investigació en el camp de la visió per computador en diversos àmbits. Es centren en la transferència de coneixement i tecnologia cap a les empreses[5].

Aquest projecte, tot i tenir una gran part tècnica i tenir com a entregable final un SIEM desplegat s'ha enfocat des d'un punt de vista de consultoria. Des de l'inici del projecte s'ha tractat al CVC com a un client al qual se li dona assessoria en matèria de ciberseguretat. S'ha depès de la informació proporcionada pel CVC per definir els requeriments del projecte i sempre s'ha actuat a mode de recomanació, amb el CVC aprovant totes les accions efectuades al seu entorn.

3.1 El CVC com a entorn tecnològic

Per a realitzar un projecte de les característiques és vital disposar de una visió global de la organització i del seu entorn tecnològic.

Per obtenir aquesta visió es van realitzar diverses reunions amb tècnics del CVC per recollir requeriments i entendre els següents conceptes:

- Què és el CVC
- Quina forma té el seu entorn tecnològic
- Fonts d'informació que podrien ser rellevant integrar al SIEM
- Dificultats tècniques que es puguin tenir

Aquestes reunions, que es van dur a terme a l'inici del projecte van servir per extreure múltiples conclusions i dibuixar el pla de ruta que es va aplicar en les posteriors etapes.

El CVC té molts investigadors externs que es connecten al seu entorn. Això provoca que els seus controls de IAM (Identity and Access Management) hagin de ser més flexibles de l'habitual.

El CVC no disposava d'un mapa de xarxa actualitzat però es va poder dibuixar un amb la informació recol·lectada. A grans trets, la informació que vam determinar com a rellevant va ser la següent:

- És una organització petita, a nivell d'infraestructura la porta poca gent.
- Tenen ordinadors corporatius (endpoints) tant amb sistemes operatius Linux com Windows, els servidors son predominantment Linux.
- Es disposa d'un tallafocs Watchguard. Per en-entrar/sortir a internet es requereix passar pel tallafocs.
- Hi ha dues solucions antivirus instal·lades als actius del CVC: un Trend Micro proporcionat per l'UAB i Windows Defender als endpoints d'usuari.
- Utilitzen servidors de correu al núvol, concretament Office365.
- Hi ha dos controladors de domini.
- Disposen de DHCP.

Un cop recollida aquesta informació inicial es va poder passar a la fase de disseny de la solució.

4 Disseny de la solució

4.1 Arquitectura

A l'hora de dissenyar la solució més adient per al CVC es van tenir en compte tots els factors comentats a les reunions de seguiment.

La primera decisió que es va prendre va ser quin SIEM implementar al CVC. Actualment hi ha més de 50 solucions SIEM al mercat, cadascuna amb les seves funcionalitats, característiques, avantatges i desavantatges.

Per triar el SIEM que millor resultats pogués donar es va tenir en compte una premissa inicial no modificable: el software ha de ser gratuït. Els SIEMs són eines que es venen com a SaaS i la falta de pressupost del projecte va descartar tots aquells que tinguessin un llicenciament no gratuït.

Es van avaluar 7 opcions al detall i al final va haver-hi dos finalistes, IBM QRadar SIEM Community Edition i AlienVault OSSIM.

Es va fer un benchmarking detallat dels dos productes on es van avaluar les capacitats dels dos aplicatius i es va valorar quins podien ser més interesants per al CVC.

	QRadar	AlienVault
Asset Discovery	✗	✓
Vulnerability Assesment	✗	✓
Intrusion Detection	✓	✓
Event Correlation	✓	✓
Threat Intelligence	✓	✓
Dedicated support	✗	✗
Analytics and data visualization	✓	✗

Figura 2. Benchmarking QRadar vs AV[6]

Inicialment es va optar per instal·lar AlienVault, la qual cosa més endavant va resultar no ser encertat. Es va arribar a instal·lar i a fer arribar-hi esdeveniments però a

l'utilitzar l'eina ens vam adonar que la dificultat per realitzar algunes operacions dintre de l'aplicació i el fet de que no es disposés de visualització de dades a temps real feia que, a llarg termini, no sigues una eina que el CVC estigués disposat a utilitzar.

Finalment, es va optar per IBM QRadar SIEM Community Edition i es va tirar endavant amb aquesta opció.

Es va dissenyar la següent arquitectura per la solució:

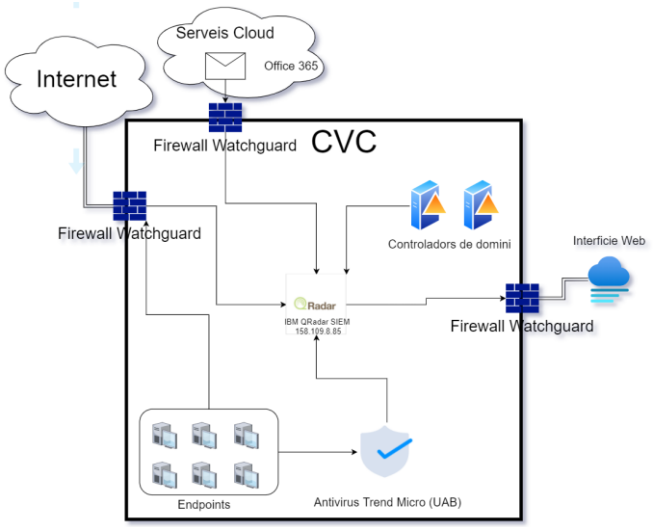


Figura 3. Arquitectura solució SIEM[7]

A la figura 3 es pot veure de manera gràfica quin és el disseny proposat per la solució.

L'element central és la consola de QRadar, a on aniran a parar els esdeveniments de seguretat de les altres eines en el diagrama.

El tallafocs Watchguard proporcionarà visibilitat perimetral de les connexions entrants i sortints de l'entorn del CVC.

L'antivirus Trend Micro és l'antivirus proporcionat per la UAB. Es va peticionar a la UAB accedir als logs de la consola però, degut a que integrant aquesta font es veurien esdeveniments de dispositius de tota la UAB i no només limitats al CVC, es va denegar l'accés als esdeveniments.

Potser heu notat l'absència del antivirus propi del CVC, el Windows Defender, al diagrama. La raó per la qual es va descartar el Windows Defender és perquè actualment el Windows Defender corre individualment a cada màquina sense tenir una consola central que orquestri. S'ha realitzat la recomanació al CVC de que centralitzin els logs en una consola central i s'ha proporcionat la documentació adient per fer-ho.

Per últim s'ha considerat interesant considerar la integració tant dels servidors de correu com dels controladors de domini. Els servidors de correu ens permetrà obtenir informació sobre els correus entrants i sortints, que és el vector més gran d'entrada de malware i el controlador de domini ens donarà visió generalitzada dels usuaris de domini i dels recursos als qual accedeixen.

6.1 Casos d'ús

En la fase de disseny vam considerar oportú fer una visita als casos d'ús.

Es defineix com a cas d'ús la descripció d'una acció o una activitat. En el nostre cas, és un escenari plantejat on es preveu la consecució d'una sèrie d'accions seqüencials que requereixen una resposta automàtica. Aquesta resposta automàtica pot ser de diferents tipus:

- Generar un nou esdeveniment
- Enviar un correu electrònic
- Notificar a un usuari o grup d'usuaris en concret
- Afegir a un Reference Set
- Executar acció personalitzada (executar un script propi o realitzar una combinació de respostes automàtiques)

De cada cas d'ús es pot crear una regla que té una resposta configurada. En el nostre cas, totes les regles tenen com a resposta automàtica notificar a l'usuari que s'han donat un seguit d'esdeveniments que val la pena revisar manualment per determinar si hi ha hagut un incident de seguretat.

Es van dissenyar casos d'ús i sub-casos d'ús específics per a cada eina que es vol integrar. Aquests casos d'ús són un avanç i es modificaran, s'eliminaran o es crearan d'altres un cop es rebin esdeveniments i es pugui realitzar un anàlisi exhaustiu d'aquests[8].

- **Controlador de domini**
  - Connexions simultànies des de diferents localitzacions geogràfiques
  - Atacs de força bruta
  - Atacs a Kerberos
  - Canvis en polítiques
  - Canvi de contrasenya d'un usuari administrador
  - Creació i eliminació d'un usuari en un breu període de temps
- **Office365**
  - Connexions simultànies des de diferents localitzacions geogràfiques
  - Atacs de força bruta
  - Detecció de mails fraudulents o *"spam"*
- **Antivirus**

- Malware detectat i no esborrat
- Un sol endpoint infectat per múltiples malwares
- Un sol malware infectant múltiples endpoints
- Infecció recurrent d'un endpoint per un mateix malware

A part dels casos d'ús personalitzats que implementarem, IBM QRadar SIEM disposa de casos d'ús predefinits que venen instal·lats per l'eina.

Tot i que aquests casos d'ús són generats per tècnics d'IBM professionals en el àmbit, al no ser personalitzats a la organització ni als esdeveniments que s'estan rebent molts cops solen donar un alt nombre de falsos positius.

5 IMPLEMENTACIÓ

Un cop finalitzada la fase de disseny cal plasmar les idees del paper en infraestructura real.

Al ser una versió gratuïta, el format en el que IBM proporciona IBM QRadar Community és l' "all-in-one", és a dir, tots els components que formen el sistema s'han d'allotjar dintre d'una sola màquina virtual. En el cas de triar la opció de llicenciament no gratuïta, es té la opció de separar els components en diferents servidors físics per a millorar la escalabilitat. La opció més popular en els últims temps és portar tota la infraestructura de QRadar a algun servei cloud però això no és possible al fer servir la versió gratuïta.

Tot i ser un desplegament "all-in-one" tots els components s'instal·len de la mateixa manera dintre d'una sola màquina virtual:

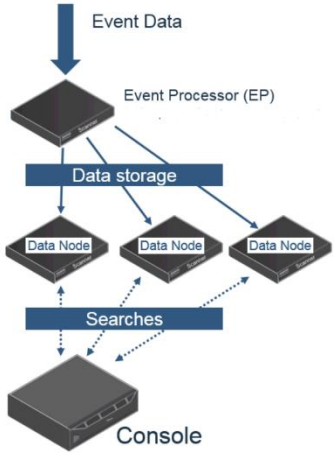


Figura 4. Fluxe de tractament de dades[9]

El primer pas per desplegar l'aplicatiu és crear la màquina virtual on s'allotjarà.

Aquesta màquina virtual ha de tenir les següents característiques mínimes[10]:

- 6GB de memòria RAM
- 250GB d'espai de disc
- 2 cores CPU
- Adaptador de xarxa amb accés a Internet
- Una IP estàtica pública i privada
- Un hostname qualificat

Per crear aquesta màquina hem usat el VMware ESXi 7.5 que té el CVC allotjat a un servidor físic.

La instal·lació es va fer de manera presencial des del propi servidor físic del CVC però, per poder-hi accedir de manera remota per realitzar tasques on és necessari accedir a la consola o tasques de recuperació en cas de que l'aplicatiu caigués es va publicar el servei web per a que pogués tenir-hi accés remot.

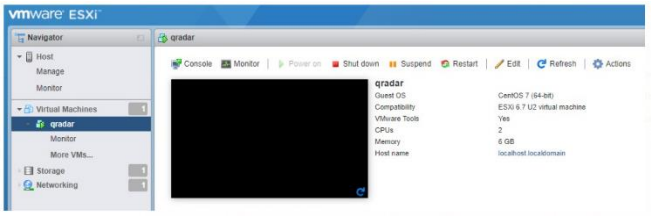


Figura 5. VMware ESXi 7.5

IBM proporciona QRadar SIEM Community Edition de manera gratuïta en format .OVA, per tant, no cal preinstal·lar cap sistema operatiu de manera prèvia.

La OVA està basada en el sistema operatiu CentOS 7.

La instal·lació es va haver de fer diversos cops degut a problemes amb la configuració de xarxa de la màquina virtual i errors de sincronització dels serveis NTP però finalment es va poder realitzar la instal·lació de manera exitosa.

Un cop finalitzada la instal·lació l'aplicatiu ens publica un servei web a la IP de gestió de la màquina virtual on podrem accedir a la interfície gràfica.

6 CONFIGURACIÓ

Un cop desplegada l'eina, podem accedir a https://<adreça_ip>/console on trobarem la interfície web de l'aplicació

Figura 6. Pantalla inici QRadar

Idealment la interfície web d'un SIEM no hauria d'estar publicada a internet, però degut a la pandèmia i la poca compatibilitat dels meus horaris i els del CVC es va decidir publicar el servei web a internet de manera temporal per permetre el treball remot.

Podem logar amb l'usuari admin establert a la instal·lació.

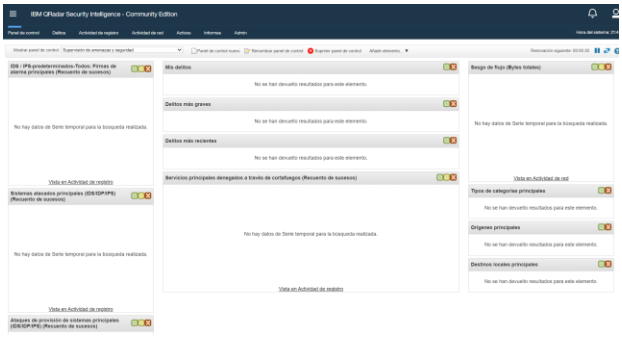


Figura 7. Dashboard QRadar

Ara mateix tenim el QRadar Community Edition desplegat i funcional, però com es pot veure a la figura 7 no es representa cap dada al tauler. Això és perquè encara no hi tenim cap dada integrada.

6.1 Integració de Dades

Per integrar les dades hem de fer arribar els esdeveniments a temps real al processador d'esdeveniments de IBM QRadar Community Edition.

IBM QRadar Community Edition escolta pel port 514 pels esdeveniments entrants.

Varis protocols de transferència de dades són vàlids per enviar informació fins a QRadar però el protocol Syslog és el que s'ha utilitzat al llarg del projecte.

A continuació es mostra l' esquema de l'enviament de logs cap al processador d'esdeveniments de IBM QRadar SIEM

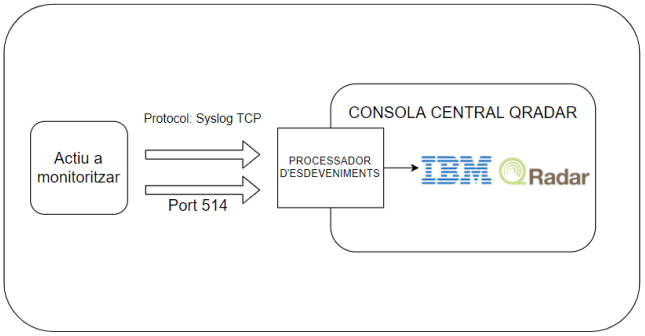


Figura 8. Diagrama d'integració d'esdeveniments

L'enviament d'esdeveniments es configura des de l'aplicatiu a monitoritzar. Pràcticament totes les eines de seguretat del mercat o qualsevol aplicatiu que corri un sistema operatiu UNIX tenen aquesta opció disponible.

És molt habitual habilitar un gestor d'esdeveniments com Logstash abans de arribar al processador d'esdeveniments per a fer un filtratge previ. En aquest cas, al ser un projecte on es preveu una ingesta de dades de volum relativament baix no ha sigut necessari.

Un cop els esdeveniments estiguin arribant correctament a la consola central de QRadar després de passar pel processador d'esdeveniments els veurem aparèixer a la finestra de "log activity".

Nombre de suceso	Origen de registro	Recuento de sucesos	Hora	Categoría de nivel bajo	IP de origen	Puerto de origen	IP de destino	Puerto de destino	Nombre de usuari
unknown	WatchGuard	1	25 ju...	Desconocido	158.109.8.9	0	158.109.8.9	0	N/A
unknown	WatchGuard	1	25 ju...	Desconocido	158.109.8.9	0	158.109.8.9	0	N/A
unknown	WatchGuard	1	25 ju...	Desconocido	158.109.8.9	0	158.109.8.9	0	N/A
unknown	WatchGuard	1	25 ju...	Desconocido	158.109.8.9	0	158.109.8.9	0	N/A
unknown	WatchGuard	1	25 ju...	Desconocido	158.109.8.9	0	158.109.8.9	0	N/A
unknown	WatchGuard	1	25 ju...	Desconocido	158.109.8.9	0	158.109.8.9	0	N/A
unknown	WatchGuard	1	25 ju...	Desconocido	158.109.8.9	0	158.109.8.9	0	N/A
unknown	WatchGuard	1	25 ju...	Desconocido	158.109.8.9	0	158.109.8.9	0	N/A
unknown	WatchGuard	1	25 ju...	Desconocido	158.109.8.9	0	158.109.8.9	0	N/A

Figura 9. Llistat d'esdeveniments a temps real

Com podem observar a la figura 9, els esdeveniments apareixen com a "unknown" i el QRadar tampoc és capaç d'identificar els altres camps com els ports d'origen i destí, l'usuari o inclús la IP d'origen o destí. Per defecte et posa la del WatchGuard i no la de la connexió que s'ha establert. Això és perquè QRadar no és capaç de identificar l'esdeveniment per sí mateix. Per a què QRadar sigui capaç d'analitzar i treballar amb els esdeveniments cal mapejar-los.

6.2 Identificació de la font, mapeig d'esdeveniments i parseig de camps clau

Un cop arriben esdeveniments al processador d'aquests, cal indi-car-li a QRadar què és el que està arribant. Inicialment, QRadar només entén que li està arribant un esdeveniment syslog. No sap quin és l'origen, ni entén el contingut del mateix.

És necessari que QRadar identifiqui correctament la font, per fer-ho, utilitza el Device-ID de la capçalera syslog.

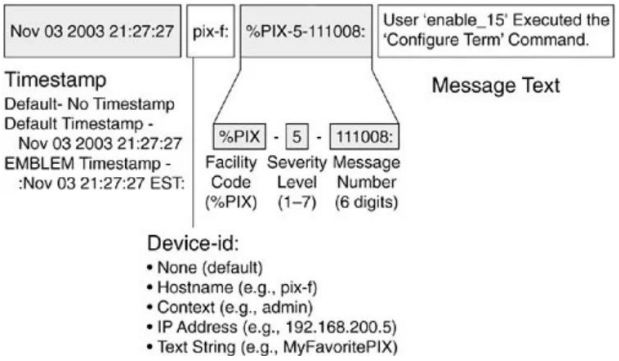


Figura 10. Estructura de la capçalera syslog[11]

En el cas de, per exemple, el tallafocs WatchGuard, el Device-ID és CVCFIRE. Indiquem a QRadar que quan entri un esdeveniment i llegeixi la capçalera, si el Device-ID és igual a CVCFIRE assigni l'esdeveniment com a originari del tallafocs WatchGuard. La configuració de la font de informació quedaria de la següent manera:

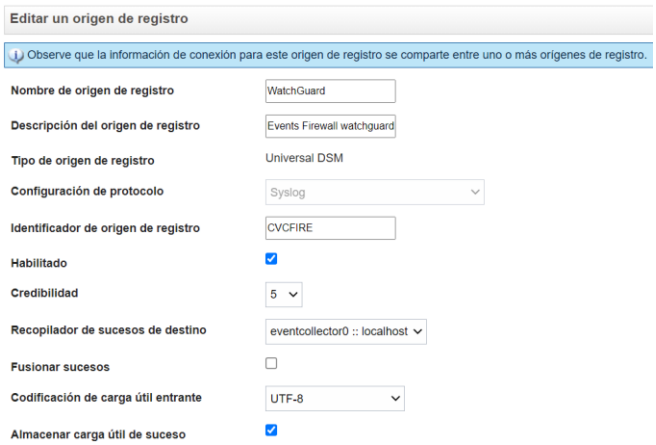


Figura 11. Editor de fonts d'informació

Un cop QRadar pot identificar la font, haurem de parsejar els camps que ens siguin necessaris. El parseig s'efectua amb expressions regulars.

Els esdeveniments presenten patrons amb determinats caràcters, amb les expressions regulars indiquem al intèrpret de QRadar que examina cada esdeveniment entrant per fer match amb l'expressió regular identificada.

Per exemple, els esdeveniments del tallafocs tenen un camp que indica la geolocalització de la connexió entrant o sortint. Aquest camp ve donat de la següent manera: "geo_src="país"". La expressió regular que s'ha definit per parsejar aquest camp és geo_src="(.*?)".

Cada vegada que un nou esdeveniment entri al processador d'esdeveniments aplicarà tots els parsers que haguem configurat per aquella font o grup de fonts en concret. Per aquest motiu és important optimitzar adequadament les expressions regulars ja que una expressió regular mal optimitzada pot alentir el processador de manera exponencial.

Per últim, és necessari indicar el nom de l'esdeveniment. El nom de l'esdeveniment és important ja que serveix per indexar els diferents tipus d'esdeveniments. Seguint amb l'exemple del tallafocs WatchGuard, els diferent tipus de noms d'esdeveniments que s'han configurat són els següents:

- Connection initiated
- Authentication
- Process start
- Process shutdown
- Task performed by device administrator

El nom de l'esdeveniment és vital per poder fer consultes precises a la consola i la posterior creació de casos d'ús, reports, cerques, regles etc.

6.3 Implementació dels casos d'ús

Un cop tenim la font identificada, els esdeveniments mapejats i els camps parsejats QRadar ja és capaç d'aplicar intel·ligència als esdeveniments i cercar per comportaments que nosaltres haguem declarat prèviament com a sospitosos.

Aquesta intel·ligència s'aplica a QRadar en la forma de regles. Les regles s'escriuen en un llenguatge propi d'IBM molt semblant al llenguatge natural basat en accions des d'un assistent o wizard.

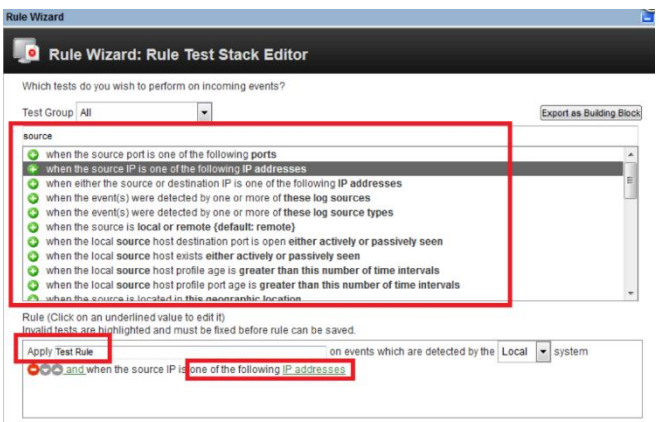


Figura 12. Editor de casos d'ús

Com podem veure a la figura 12, es poden seleccionar un seguit d'accions concatenades amb els operadors lògics AND o OR.

A mode d'exemple, si volguéssim crear una regla quan es rep un possible atac de força bruta, la lògica seria la següent:

Apply "Atac de força bruta" on events which are detected by the local System AND when event name is "Authentication Failed" AND more than 30 events are seen in 5 minutes.

Aquí veiem la importància de mapejar correctament els esdeveniments i parsejar tots els camps que siguin rellevants a nivell de seguretat. És necessari separar totes les dades per a que el sistema pugui treballar amb elles de manera correcta.

6.4 Configuració de Reports

Els reports funcionen de manera molt similar a les cerques o consultes.

Les consultes a la base de dades d'esdeveniments es realitzen amb el llenguatge AQL (Ariel Query Language) dissenyat per IBM.

Podem generar una consulta AQL i configurar el sistema per a fer-la recurrentment, per exemple cada 24 hores. Podem demanar que els resultats d'aquesta consulta se'ns enviïn per correu.

6.5 Visualització de dades

De la mateixa manera que generem reports, també podem guardar consultes en AQL que podem mostrar al panell principal a temps real.

Per una ràpida visualització de dades que es consideren critiques per la organització es poden configurar al panell principal consultes que s'executaran recurrentment per mostrar els resultats a temps real. Per exemple, aquests serien els països d'origen de les connexions entrants que passen pel tallafoc WatchGuard en la última hora.

En aquest cas, veiem que Rússia significa el 42% dels orígens de les connexions entrants al CVC, de per sí i sense més context, valdria la pena investigar més per veure el motiu d'aquestes connexions.

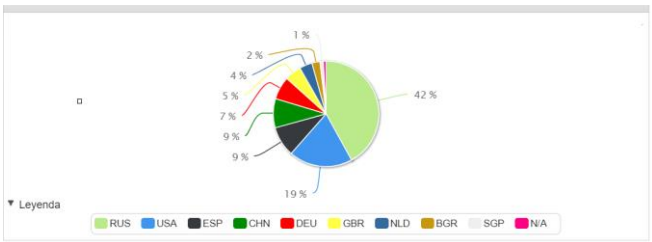


Figura 13. Gràfica d'orígens de connexions entrants

Val la pena esmentar que gran part d'aquestes connexions són denegades per el WatchGuard.

7 Resultats, següents passos i conclusions

7.1 Resultat final

Tal i com el títol indica, l'objectiu del projecte era el disseny, la implementació i la configuració d'un SIEM. Considerem que aquests objectius han estat assolits de manera clara i amb evidencies tan tangibles com que actualment el CVC disposa d'un SIEM al seu entorn a on hi arriba informació d'actius d'alta importància per a ells que es disseccciona, s'analitza i es correla usant regles lògiques i intel·ligència específicament creades per a ells.

Tot i així, això és només el començament. Un SIEM és una eina que necessita optimització constant i millora continua. S'ha de ser molt crític amb l'estat actual d'un SIEM i s'ha de treballar constantment en integrar noves fonts, crear nous casos d'ús rellevant per la organització i el què és més important, afinar els casos d'ús existents.

Per la natura de les infraestructures de TI de les organitzacions, el nombre de falsos positius i alertes de potencials incidents de seguretat que després resulten no ser tals al ser un comportament normal de la xarxa o errors de configuració és molt elevat, especialment al desplegar per primera vegada un software d'aquest tipus.

Per la nostra condició d'externs al CVC i al no tenir accés de cap tipus a la organització més enllà de accés a les màquines on s'han allotjat els aplicatius desplegats ens és gairebé impossible realitzar investigacions sobre les alertes generades. Totes les alertes que s'han generat, s'han enviat a administradors del CVC per que les puguin revisar però cal dir que creiem que aquestes alertes són de poca qualitat; hi ha un nombre molt elevat d'elles i, sense saber-ho amb exactitud, preveiem que el nombre de falsos positius sigui molt elevat.

En un món ideal el nombre d'alertes que sorgeixen a diari és petit i la credibilitat d'aquestes és molt elevada. Com a consultor en matèria de ciberseguretat, he participat en molts projectes d'optimització d'un SIEM on s'han dedicat centenars, si no milers d'hores, per arribar a aquest nivell de maduresa. Aquesta feina d'optimització i millora continua no estava dintre de l'abast del projecte.

Actualment, el CVC disposa d'un anàlisi del seu entorn des d'un punt de vista de ciberseguretat, el disseny d'una arquitectura base optima per extreure el màxim de les capacitats del SIEM, d'un SIEM a on hi arriba informació rellevant a nivell de seguretat que es mapeja i parseja per a un tractament posterior.

7.2 Següents Passos

Arribats a la finalització del projecte, amb un SIEM funcional desplegat, està en mans del CVC decidir què volen fer amb ell i com afronten tots els reptes que comporta operar-lo.

Cal dir que el CVC em va fer saber que no disposaven del temps ni dels recursos necessaris per operar una eina tant demandant com és un SIEM. Tal i com vam recollir a les reunions de recollida de requeriments, el CVC és un entorn petit que el porta poca gent i, de per sí, estan prou enfeinats amb els seus propis projectes. Tenir un recurs dedicat a la ciberseguretat i invertir totes les hores que un SIEM necessita no tindria sentit per una organització com el CVC. Tot i així, es poden donar altres usos a un SIEM.

Es cert que un SIEM és una eina que s'ha dissenyat específicament per ajudar als equips de TI a orquestrar les seves operacions de seguretat i a tenir una visibilitat més elevada del seu entorn tecnològic però les capacitats d'un SIEM es poden explotar de maneres més creatives i no relacionades amb la seguretat.

La raó per la qual la visualització de dades en temps real era tant important per nosaltres i per al CVC és perquè en les reunions de seguiment ens van fer saber que una utilitat que ells li podrien donar al SIEM és enviar-li esdeveniments de aplicacions pròpies no relacionats amb la ciberseguretat per ajudar-los a fer un tractament de dades d' aquests.

El fet de poder mapejar els esdeveniments, parsejar camps claus, indexar-los i realitzar consultes sobre els mateixos els hi podria ser útil. Si a això li sumes la capacitat d'aplicar regles lògiques per detectar heurístiques concretes i realitzar accions automàtiques quan es donen certs escenaris converteixen a un SIEM en un agregador de dades molt potent. Tot i que no és el propòsit amb el que s'ha implantat la solució, al cap i a la fi, tant el servidor on està allotjat el software com la informació que s'hi envia són propietat del CVC i poden donar-li l'ús que els hi sembli més adient.

Com a expert específicament en IBM QRadar i com no podria ser d'un altre manera, m'he ofert als administradors del CVC a donar formacions, fer demostracions i donar tot el suport que sigui necessari durant un període de temps il·limitat per a que puguin exprimir al màxim totes les funcionalitats que un SIEM els hi pot donar.

7.3 Conclusions

Realitzar aquest projecte ha sigut una gran experiència per mi. Tot i que he treballat durant gairebé 4 anys fent projectes molt similars a grans empreses de l'àrea de Barcelona aquest projecte ha estat diferent per diversos motius.

En primer lloc, no he tingut el suport tècnic d'una gran empresa darrere meu. Tot i liderar jo els projectes, la meva empresa disposava dels millors tècnics en ciberseguretat de Catalunya a on podia acudir si tenia dubtes o les coses no sortien bé a la primera. En el cas d'aquest TFG, tot i que és innegable que he tingut l'ajuda del meu tutor en tot moment, a nivell tècnic he hagut de desenvolupar-me sol i això ha estat un gran aprenentatge.

En segon lloc, ha sigut fascinant realitzar aquest projecte en una organització com el CVC. En els projectes d'aquest tipus que he realitzat a altres empreses, hi havia un nivell de maduresa a nivell de ciberseguretat molt més elevat. El SIEM només havia de complementar les seves operacions de seguretat. En el cas del CVC això no ha estat així. A nivell de seguretat no existeix una preocupació real i prèviament no es realitza cap operació de seguretat més enllà de les típiques d'un administrador de sistemes. Veure aquesta altra cara ha sigut molt enriquidor.

Aquest últim punt ha fet que veiés coses que no son típiques en una empresa més madura tecnològicament, i ha fet que el nombre d'alertes generades fos molt elevat.

El nivell de maduresa global en ciberseguretat del CVC és baix, això fa que hi hagi especialment molta feina a l'hora d'operar el SIEM i tenen molts reptes per davant i molt marge de millora.

Agraïments

Agraïments a la Marta per suportar-me, a l'Andreu pel seu esplèndid treball tutoritzant el projecte i, sobretot, al CVC per dedicar-me hores del seu temps que no tenien i cedir-me les seves instal·lacions i infraestructura per portar a terme el projecte.

Bibliografia

[1]**Que és un SIEM i com funciona.** *https://www.fireeye.com/products/helix/what-is-siem-and-how-does-it-work.html*

[2]**Cybersecurity Market Report.** *https://cybersecurityventures.com/cybersecurity-market-report/*

[3]**Cybersecurity Market Trends.** *https://www.cyber-observer.com/cyber-news-29-statistics-for-2020-cyber-observer/*

[4]**Video: SoC Structure and components.** *https://www.f-secure.com/en/consulting/our-thinking/purple-teaming/people-process-technology*

[5]**Pàgina oficial Centre de Visió per Computador.** *http://www.cvc.uab.es/?lang=ca*

[6]**Benchmarking de SIEMs** */Recursos/Benchmarking.xlxs*

[7]**Arquitectura de la solució.** */Recursos/Arquitecture.drawio*

[8]**Casos d'ús.** */Recursos/casos-d-us.docx*

[9]**IBM QRadar Data Processing Flow.** *https://www.ibm.com/support/pages/qradar-configuring-16xx18xx-appliances-processing-only-mode*

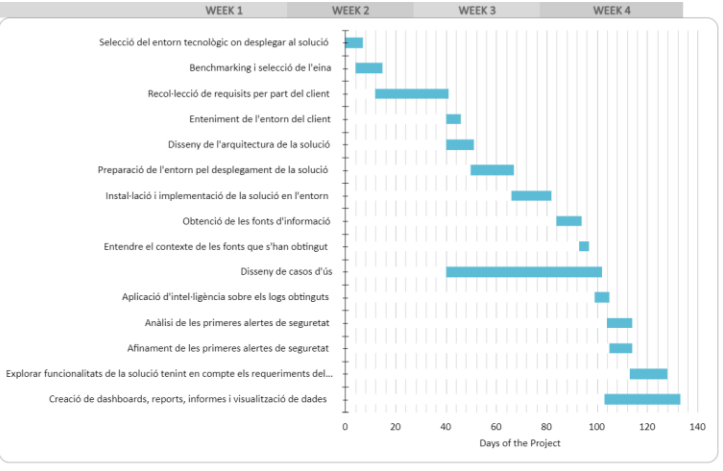
[10]**Documentació oficial IBM QRadar Community Edition 7.3.3.** */Documentacio/b_qradar_community_edition_7.3.3GA_v1.0.pdf*

[11]**Estructura i format dels missatges Syslog.** *https://stackify.com/syslog-101/*

APÈNDIX

A1. PLANIFICACIÓ DEL TREBALL

TASK NAME	START DATE	END DATE	START ON DAY*	DURATION* (WORK DAYS)	PERCENT COMPLETE
Desplegament d'un SIEM al CVC					
Selecció del entorn tecnològic on desplegar al solució	2/8	2/14	0	7	100%
Benchmarking i selecció de l'eina	2/12	2/22	4	11	100%
Recol·lecció de requisits per part del client	2/20	3/20	12	29	100%
Enteniment de l'entorn del client	3/20	3/25	40	6	100%
Disseny de l'arquitectura de la solució	3/20	3/30	40	11	100%
Preparació de l'entorn pel desplegament de la solució	3/30	4/15	50	17	100%
Instal·lació i implementació de la solució en l'entorn	4/15	4/30	66	16	100%
Obtenció de les fonts d'informació	5/3	5/12	84	10	100%
Entendre el contexte de les fonts que s'han obtingut	5/12	5/15	93	4	100%
Disseny de casos d'ús	3/20	5/20	40	62	100%
Aplicació d'intel·ligència sobre els logs obtinguts	5/18	5/23	99	6	100%
Anàlisi de les primeres alertes de seguretat	5/23	6/1	104	10	100%
Afinament de les primeres alertes de seguretat	5/24	6/1	105	9	100%
Explorar funcionalitats de la solució tenint en compte els requeriments del client	6/1	6/15	113	15	100%
Creació de dashboards, reports, informes i visualització de dades	5/22	6/20	103	30	100%



A2. OPERACIONS LOGIQUES DE REGEX

Anchors

^	Start of string, or start of line in multi-line pattern
\A	Start of string
$	End of string, or end of line in multi-line pattern
\Z	End of string
\b	Word boundary
\B	Not word boundary
\<	Start of word
\>	End of word

Character Classes

\c	Control character
\s	White space
\S	Not white space
\d	Digit
\D	Not digit
\w	Word
\W	Not word
\x	Hexadecimal digit
\O	Octal digit

POSIX

[:upper:]	Upper case letters
[:lower:]	Lower case letters
[:alpha:]	All letters
[:alnum:]	Digits and letters
[:digit:]	Digits
[:xdigit:]	Hexadecimal digits
[:punct:]	Punctuation
[:blank:]	Space and tab
[:space:]	Blank characters
[:cntrl:]	Control characters
[:graph:]	Printed characters
[:print:]	Printed characters and spaces
[:word:]	Digits, letters and underscore

Quantifiers

*	0 or more	{3}	Exactly 3
+	1 or more	{3,}	3 or more
?	0 or 1	{3,5}	3, 4 or 5

Add a ? to a quantifier to make it ungreedy.

Escape Sequences

\	Escape following character
\Q	Begin literal sequence
\E	End literal sequence

"Escaping" is a way of treating characters which have a special meaning in regular expressions literally, rather than as special characters.

Common Metacharacters

^	[	.	$
{	*	(	\
+	)	|	?
<	>		

The escape character is usually \

Special Characters

\n	New line
\r	Carriage return
\t	Tab
\v	Vertical tab
\f	Form feed
\xxx	Octal character xxx
\xhh	Hex character hh

Groups and Ranges

.	Any character except new line (\n)
(a|b)	a or b
(...)	Group
(?:...)	Passive (non-capturing) group
[abc]	Range (a or b or c)
[^abc]	Not (a or b or c)
[a-q]	Lower case letter from a to q
[A-Q]	Upper case letter from A to Q
[0-7]	Digit from 0 to 7
\x	Group/subpattern number "x"

Ranges are inclusive.

Pattern Modifiers

g	Global match
i *	Case-insensitive
m *	Multiple lines
s *	Treat string as single line
x *	Allow comments and whitespace in pattern
e *	Evaluate replacement
U *	Ungreedy pattern

* PCRE modifier

String Replacement

$n	nth non-passive group
$2	"xyz" in /^(abc(xyz))$/
$1	"xyz" in /^(?:abc)(xyz)$/
$`	Before matched string
$'	After matched string
$+	Last matched string
$&	Entire matched string

Some regex implementations use \ instead of $.